Toward an automatic method of modal analysis

Piero Barone

Istituto per le Applicazioni del Calcolo "M. Picone", C.N.R.

via dei Taurini 19, 00185 Rome, Italy

E-mail: piero.barone@gmail.com; p.barone@iac.cnr.it

Abstract. A common problem, arising in many different applied contexts, consists in estimating the number of exponentially damped sinusoids whose weighted sum best fits a finite set of noisy data and in estimating their parameters. Many different methods exist to this purpose. The best of them are based on approximate Maximum Likelihood estimators, assuming to know the number of damped sinusoids, which is then estimated by an order selection procedure. It turns out that Maximum Likelihood estimators are biased in this specific case. The idea pursued here is to cope with the bias, by a stochastic perturbation method, in order to get an estimator with smaller Mean Squared Error than the Maximum Likelihood one. Moreover the problem of estimating the number of damped sinusoids and the problem of estimating their parameters are solved jointly. The method is automatic, provided that a few hyperparameters have been chosen, and faster than standard best alternatives.

MSC: 30E10, 65C60

Keywords: complex moments problem, random Hankel pencils, stochastic perturbations

Introduction

Let's consider the model

$$f_R(t;q,P_R)=\sum_{j=1}^{q} A_j\rho_j^t\cos(2\pi\omega_j t+\theta_j),\ \ t\in I\!R^+,\ \ \omega_h\neq\omega_k\ \forall h,k, \tag{1}$$

$$P_R=\{A_j,\rho_j,\omega_j,\theta_j,\ j=1\ldots,q\}\in I\!R^{4q} \tag{2}$$

and assume that we want to estimate q, P_R from the data

$$a_k=f_R(k\Delta)+\epsilon_k,\ k=0,\ldots,n-1,n\geq 4q,\Delta\in R^+$$

where $\Delta>0$ is known, ϵ_k are i.i.d. zero-mean Gaussian variables with known variance σ^2. In order to make the model f_R identifiable from $\{a_k\}$ we assume that $|\omega_j|\Delta\leq\pi,\ \forall j$. In fact if e.g. $\omega_r\Delta>\pi$ it exists $\tilde\omega\in[-\pi,\pi]$ such that $\omega_r\Delta=\tilde\omega\Delta+2\pi h, h\in I\!N, h\neq 0$ and $f_R(t;q,P_R)=f_R(t;q,P_R')$ where $P_R'=P_R\setminus\{\omega_r\}\bigcup\{\tilde\omega\}$. We notice that $f_R(t,q,P_R)$ is a particular case of the complex model

$$f(t;p,P)=\sum_{j=1}^{p}c_j\xi_j^t,\ \ t\in I\!R^+,$$
$$P=\{c_j,\xi_j,\ j=1,\ldots,p\}\in \mathbb{C}^{2p}$$

when $p=2q, q\in I\!N,\ \Im m(f)=0$ and

$$c_j=\frac{1}{2}A_je^{i\theta_j},\ \xi_j=\rho_je^{i2\pi\omega_j}, j=1,\ldots,q,$$
$$c_j=\frac{1}{2}A_{j-q}e^{-i\theta_{j-q}},\ \xi_j=\rho_{j-q}e^{-i2\pi\omega_{j-q}}, j=q+1,\ldots,p.$$

Therefore in the following we consider the problem of estimating P from the complex data $(a_k,\ k=0,\ldots,n-1)$ with the identifiability condition $|arg(\xi_j)|\Delta\leq\pi\ \forall j$, where the noise ϵ_k are i.i.d. zero-mean complex Gaussian variables with known variance σ^2 i.e. the real and imaginary parts of a_k are independently distributed as Gaussian variables with variance $\sigma^2/2$ and mean $\Re e[f(k\Delta)],\Im m[f(k\Delta)]$ respectively.

The problem described above arises in many fields. A certainly not exhaustive list is the following: noisy Hausdorff moment problem, numerical inversion of Laplace transform, noisy trigonometric moment problem, identification of constant coefficients

ODE from its transient response, approximation by complex exponentials functions, modal analysis, direction of arrival problem, shape from moments problem [6, 7, 8, 13, 16, 18, 25, 27].

It is well known that the problem can be severely ill posed, depending on the relative location in the complex plane of the points $\xi_j, j = 1, \ldots, p$ and on the ratios $SNR_j = |c_j|/\sigma, j = 1, \ldots, p$. A further difficulty is related to the fact that p is unknown. This means that when the ratios $SNR_j, j = 1, \ldots, p$ are bounded by some constant $C < \infty$ even if you are able to guess the right order p of the model, different realization of the process a_k can give rise to quite different estimates of the other parameters in P. The difficulty of guessing the right order is related to the difficulty of estimating the other parameters. In fact if these were correctly estimated a good guess of p would minimize an order selection criterium such as AIC or BIC [2]. Unfortunately you cannot hope to get good estimates of the other parameters if p is not correctly estimated. Because of this situation many methods have been proposed to solve the problem by filtering the noise in different ways and/or considering different estimators. Those which provide the best performances, assuming to know the right order p, compute an approximation of the Maximum Likelihood estimator of the parameters filtering somewhat the noise at the same time [18, 19, 20]. The guess of the order is then used to build the noise filter and therefore to improve the estimates of the other parameters. Different guesses can then be tested in order to minimize e.g. BIC. An automatic procedure can then be devised. However it turns out that MLEs are biased. Therefore approximating them is not necessarily such a good idea especially in low SNRs cases. Moreover the methods which perform better need to solve one or more eigenvalues problems which can be very computationally expensive for large data sets.

In this paper a black-box method which encompasses all these difficulties is proposed and experimentally compared with one of the best known standard method (GPOF [18]) coupled with an order selection criterium (BIC). It exploits some recent results given in [3] where the bias of MLEs is used to improve the MSE by a stochastic perturbation approach as well as a method to estimate the distribution in the complex plane of the $\xi_j, j = 1, \ldots, p$ which are the most critical parameters. The problem is stated in a stochastic framework where the estimation of all the parameters in P can be addressed jointly. Moreover the method is fast and can be easily specialized to take into account prior information on the problem at hand.

The paper is organized as follows. In section 1 the Maximum Likelihood and related estimators and their properties in this context are shortly reviewed. In section 2 the proposed method is described and some parameters required to make it automatic are discussed and estimated. In section 3 a simulation experiment is presented to compare the results provided by a standard method and the proposed one.

1. Maximum Likelihood and related estimates

1.1. Algebraic and statistical properties of MLE

Maximum likelihood estimates P_{ML} of the parameters P of the model $f(t; p, P)$, assuming that p is known, are obtained by

$$P_{ML} = \mathrm{argmax}_P \, e^{-\frac{\|\underline{a} - f(\underline{t}; p, P)\|_2^2}{\sigma^2}} = \mathrm{argmin}_P \, \|\underline{a} - f(\underline{t}; p, P)\|_2^2$$

where $\underline{a} = [a_0, \ldots, a_{n-1}]$, $\underline{t} = [0, \Delta, \ldots, (n-1)\Delta]$. In order to solve this nonlinear least squares problem, following [14], we notice that the problem is separable. In fact we can split the parameters P in two sets $P = P_c \bigcup P_\xi$ where $f(t; p, \underline{\gamma}, \underline{\zeta}) = \sum_{j=1}^{p} \gamma_j \zeta_j^t$. For each fixed value $\underline{\zeta} \in P_\xi$ let us consider the function $\underline{\gamma}(\underline{\zeta})$ defined by

$$\begin{aligned}\underline{\gamma}(\underline{\zeta}) &= \mathrm{argmin}_{\underline{\gamma}} \|\underline{a} - f(\underline{t}; p, \underline{\gamma}, \underline{\zeta})\|_2^2 = \mathrm{argmin}_{\underline{\gamma}} (\underline{a} - V\underline{\gamma})^H (\underline{a} - V\underline{\gamma}) \\ &= (V^H V)^{-1} V^H \underline{a}\end{aligned}$$

where $V = V(\underline{\zeta})$ is the Vandermonde matrix of order $n \times p$ of the vector $\underline{\zeta}$, H denotes transposition plus conjugation and I_n is the identity matrix of order n. It is proved in [14] that, substituting $\underline{\gamma}(\underline{\zeta})$ in $\|\underline{a} - f(\underline{t}; p, \underline{\gamma}, \underline{\zeta})\|_2^2$ and minimizing w.r.to $\underline{\zeta}$, we get

$$\begin{aligned}
&\underline{\xi}_{ML} = \text{argmin}_{\underline{\zeta}}\|\underline{a} - f(\underline{t}; p, \underline{\gamma}(\underline{\zeta}), \underline{\zeta})\|_2^2 = \\
&\text{argmin}_{\underline{\zeta}}(\underline{a} - V(V^H V)^{-1}V^H\underline{a})^H(\underline{a} - V(V^H V)^{-1}V^H\underline{a}) = \\
&\text{argmin}_{\underline{\zeta}}\underline{a}^H(I_n - V(V^H V)^{-1}V^H)\underline{a}
\end{aligned}$$

and

$$\underline{c}_{ML} = \underline{\gamma}(\underline{\xi}_{ML}).$$

In order to study the properties of the ML estimator we start by noticing that

Proposition 1 *It does not exist an efficient estimator of the parameters P. Specifically the MLE of P is not efficient.*

<u>Proof.</u> We notice that the log-likelihood function is an absolute continuous function of P. Hence, by Corollary 3.1 and Theorem 3.1 of [21] if the variance of an estimator of P would attain the Cramer-Rao bound this would imply that the probability density

$$\frac{1}{(\pi\sigma^2)^n} e^{-\frac{\|\underline{a} - f(\underline{t}; p, P)\|_2^2}{\sigma^2}}$$

of $\underline{a}$ would belong to the exponential family. But this is false because of the dependence of a_h on ξ^h. □

1.2. Approximate MLE: complex exponentials interpolation

We then consider the problem of interpolating the data $\underline{a}$ by means of a linear combination of complex exponential functions $\tilde{\zeta}_j^t$, $\tilde{\zeta}_j \in \mathbb{C}$, $j = 1, \ldots, n/2$, that is to find n complex numbers $[\underline{\tilde{\gamma}}, \underline{\tilde{\zeta}}] = \{\tilde{\gamma}_j, \tilde{\zeta}_j\}, j = 1, \ldots, n/2$ such that $\underline{a} = V(\underline{\tilde{\zeta}})\underline{\tilde{\gamma}}$. Equivalently we could consider the problem of building the Pade' approximation $[n/2, n/2-1]$ to the $Z-$transform of $a_k, k = 0, 1, \ldots$ [16, 7]. To this aim let us consider the Hankel matrix pencil $U_1 - zU_0, \;\; z \in \mathbb{C}$ where

$$U_0(\underline{a}) = U(a_0, \ldots, a_{n-2}), \quad U_1(\underline{a}) = U(a_1, \ldots, a_{n-1})$$

and

$$U(x_1, \ldots, x_{n-1}) = \begin{bmatrix} x_1 & x_2 & \ldots & x_{n/2} \\ x_2 & x_3 & \ldots & x_{n/2+1} \\ \cdot & \cdot & \ldots & \cdot \\ x_{n/2} & x_{n/2+1} & \ldots & x_{n-1} \end{bmatrix}$$

It is well known (e.g.[16]) that, provided that $detU_0 \neq 0, detU_1 \neq 0$, a unique solution exists which is given by $\underline{\tilde{\zeta}} = \underline{\xi}_{GE}$, where $\underline{\xi}_{GE}$ are the generalized eigenvalues of the pencil $U_1 - zU_0$ and $\underline{\tilde{\gamma}} = W_{GE}^T \underline{a}$ where W_{GE} is the matrix of generalized eigenvectors of $U_1 - zU_0$ and T denotes transposition. Moreover it turns out that $W_{GE} = \tilde{V}(\underline{\xi}_{GE})^{-T}$ where $\tilde{V}(\underline{\xi}_{GE})$ is the square Vandermonde matrix based on $\underline{\xi}_{GE}$. These properties can be easily checked by noticing that if $\underline{a} = V(\underline{\tilde{\zeta}})\underline{\tilde{\gamma}}$ then

$$U_0 = \tilde{V}(\underline{\tilde{\zeta}}) C \tilde{V}(\underline{\tilde{\zeta}})^T, \;\; U_1 = \tilde{V}(\underline{\tilde{\zeta}}) C Z \tilde{V}(\underline{\tilde{\zeta}})^T$$

where

$$C = diag\{\tilde{\gamma}_1, \ldots, \tilde{\gamma}_{n/2}\} \text{ and } Z = diag\{\tilde{\zeta}_1, \ldots, \tilde{\zeta}_{n/2}\}$$

and therefore $U_1 \tilde{V}(\underline{\tilde{\zeta}})^{-T} = U_0 \tilde{V}(\underline{\tilde{\zeta}})^{-T} Z$ which implies that $\underline{\tilde{\zeta}}$ are the generalized eigenvalues of the pencil $U_1 - zU_0$. The relation between $[\underline{c}_{ML}, \underline{\xi}_{ML}]$ and $[\underline{\tilde{\gamma}}, \underline{\tilde{\zeta}}]$ is given by

Proposition 2 *If $n = 2p$ then $[\underline{c}_{ML}, \underline{\xi}_{ML}] = [\underline{\tilde{\gamma}}, \underline{\tilde{\zeta}}]$.*

Proof. Let be $V = V(\underline{\tilde{\zeta}})$. Substituting $\underline{a} = V\underline{\tilde{\gamma}}$ in $\underline{a}^H(I_n - V(V^H V)^{-1} V^H)\underline{a}$ we get

$$\underline{\tilde{\gamma}}^H V^H (I_n - V(V^H V)^{-1} V^H) V \underline{\tilde{\gamma}} = 0.$$

But $\underline{a}^H(I_n - V(V^H V)^{-1} V^H)\underline{a} \geq 0$, hence $\|\underline{a} - f(\underline{t}; p, \underline{\gamma}(\underline{\zeta}), \underline{\zeta})\|_2^2$ takes its least possible value when $V = V(\underline{\tilde{\zeta}})$ therefore $\underline{\tilde{\zeta}} = \underline{\xi}_{ML}$ and $\underline{\tilde{\gamma}} = (V^H V)^{-1} V^H \underline{a} = \underline{c}_{ML}$. □

1.3. Bias of MLE

As an easy consequence of the Proposition above we show that the MLE can not have moments. In particular MLE can not have the mean, therefore bias can not be defined. Let us consider the case when $n = 2, p = 1, \theta_1 = \omega_1 = 0, |\rho| < 1$. Therefore

$$a_0 = A + \epsilon_0,\ a_1 = A\rho + \epsilon_1,\ \ U_0 = a_0,\ U_1 = a_1,\ \ \rho_{ML} = \frac{a_1}{a_0}.$$

The density of ρ_{ML} is then the density of the ratio of two independent Normal variables with means A and $A\rho$ respectively and variance σ^2. Assuming for simplicity that $A = 1$, it can be shown (e.g.[11] and references therein) that this density is given by

$$p_2(x) = \frac{1}{\pi(x^2+1)}\left(e^{-\frac{\rho^2+1}{2\sigma^2}} + \frac{\sqrt{2\pi}(1+\rho x)}{2\sigma\sqrt{x^2+1}}\mathrm{Erf}\left[\frac{1+\rho x}{\sigma\sqrt{2(x^2+1)}}\right] e^{-\frac{(\rho-x)^2}{2\sigma^2(x^2+1)}}\right)$$

which is a Cauchy-like density and therefore moments do not exist. However in the general case we can consider the formal power series of $\underline{\xi}_{ML}$, truncate it e.g. after the first term and compute the expectation, denoted, with abuse of notation, by $E[\underline{\xi}_{ML}]$. If we define as bias the quantity $\|E[\underline{\xi}_{ML}] - \underline{\xi}\|$ we have

Proposition 3 *When $n = 2p$ the MLE $[\underline{c}_{ML}, \underline{\xi}_{ML}]$ are biased.*

<u>Proof.</u> If $n = 2p$ by Proposition 2 $[\underline{c}_{ML}, \underline{\xi}_{ML}] = [\underline{\tilde{\gamma}}, \underline{\tilde{\zeta}}]$ which solve the complex exponentials interpolation problem. However in [3, lemma2] it was proved that $E[\underline{\tilde{\gamma}}] - \underline{c} = o(\sigma)$ and $E[\underline{\tilde{\zeta}}] - \underline{\xi} = o(\sigma)$. $\square$

Therefore when $n = 2p$ the MLE are only asymptotically unbiased in the limit for $\sigma \to 0$ in a weak sense i.e. when expectations are computed w.r.t an approximation of the true density of the ML estimator. One could argue that letting $n \to \infty$ for p fixed could improve the situation when $\sigma > 0$. This is not the case as we now show for the simplest case of the model $a_t = \rho^{(t-1)} + \epsilon_t,\ \ t = 0, \ldots, n-1$ where $|\rho| < 1$ and ϵ_t are i.i.d. Gaussian zero-mean random variables with variance σ^2. We have

Proposition 4 *The density of the MLE of the parameter ρ can be approximated by a density $p_n(x)$ such that*

$$\lim_{n\to\infty} p_n(x) = 0 \ \textit{if}\ |x| \geq 1$$

$$\lim_{n\to\infty} p_n(x) = p_\infty(x), \ x \in (-1,1)$$

and $p_\infty(x)$ is a density such that

$$\lim_{\sigma\to 0} p_\infty(x;\rho,\sigma) = \delta(x-\rho)$$

(in the sense of distributions). Moreover, for $\sigma > 0$, $p_\infty(x)$ is bimodal, the modes $\mu_{1,2}$ have opposite signs and $\lim_{\sigma\to\infty}\mu_{1,2} = \pm 1$.

Proof. Because ρ_{ML} annihilates the gradient of the likelihood which is Gaussian, following [1], we can approximate the density of ρ_{ML} for any n by considering a first order Taylor series approximation of ρ_{ML} around the point in which we want to approximate its density. We then get

$$p_n(x) = \frac{1}{K_n}\frac{1}{\sqrt{2\pi\sigma}^2}\sqrt{\sum_{j=1}^{n-1} j^2 x^{2(j-1)}}\, e^{-\frac{\left(\sum_{j=1}^{n-1} j x^{(j-1)}(x^j-\rho^j)\right)^2}{2\sigma^2\sum_{j=1}^{n-1} j^2 x^{2(j-1)}}}$$

and

$$K_n = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\sqrt{\sum_{j=1}^{n-1} j^2 x^{2(j-1)}}\, e^{-\frac{\left(\sum_{j=1}^{n-1} j x^{(j-1)}(x^j-\rho^j)\right)^2}{2\sigma^2\sum_{j=1}^{n-1} j^2 x^{2(j-1)}}}\, dx$$

which exists because

$$\lim_{x\to\pm\infty} x^2\sqrt{\sum_{j=1}^{n-1} j^2 x^{2(j-1)}}\, e^{-\frac{\left(\sum_{j=1}^{n-1} j x^{(j-1)}(x^j-\rho^j)\right)^2}{2\sigma^2\sum_{j=1}^{n-1} j^2 x^{2(j-1)}}} = 0, \quad 0 < \sigma < C < \infty.$$

It then follows that

$$\lim_{n\to\infty} p_n(x) = 0 \text{ if } |x| \geq 1$$

and

$$p_\infty(x) = \lim_{n\to\infty} p_n(x) = \frac{1}{K_\infty}\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\rho)^2}{2\sigma^2}R(x;\rho)}\sqrt{S(x)}$$

where

$$S(x) = -\frac{1+x^2}{(x-1)^3(x+1)^3}, \quad R(x;\rho) = \frac{(\rho x^3-1)^2}{(1-x)(1+x)(\rho x-1)^4(1+x^2)}$$

and

$$K_\infty = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-1}^{1} e^{-\frac{(x-\rho)^2}{2\sigma^2}R(x;\rho)} \sqrt{S(x)}dx$$

exists and it is finite because

$$\lim_{x\to 1^-} (x-1)^{\frac{1}{2}} e^{-\frac{(x-\rho)^2}{2\sigma^2}R(x;\rho)} \sqrt{S(x)} = 0, \quad 0 < \sigma < C < \infty$$

$$\lim_{x\to -1^+} (x+1)^{\frac{1}{2}} e^{-\frac{(x-\rho)^2}{2\sigma^2}R(x;\rho)} \sqrt{S(x)} = 0, \quad 0 < \sigma < C < \infty.$$

As $R(x;\rho)$ has no real zeros in $(-1,1)$, we have

$$\lim_{\sigma\to 0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\rho)^2}{2\sigma^2}R(x;\rho)} = \begin{cases} 0 & \forall x \neq \rho, \ x \in (-1,1) \\ \infty & x = \rho \end{cases}$$

therefore

$$\lim_{\sigma\to 0} p_\infty(x;\rho,\sigma) = \delta(x-\rho)$$

in the weak sense. To complete the proof let us consider the function

$$f(x;\rho,\sigma) = \frac{1}{2}\log S(x) - \log\sigma - \frac{(x-\rho)^2}{2\sigma^2}R(x,\rho) = cost. + \log[p_\infty(x)].$$

Its first derivative is

$$f'(x;\rho,\sigma) = \frac{P_1(x;\rho)}{\sigma^2 Q_1(x;\rho)} + \frac{P_2(x;\rho)}{Q_2(x;\rho)}$$

where $P_1(x;\rho), Q_1(x;\rho), P_2(x;\rho), Q_2(x;\rho)$ are polynomials of orders $15, 17, 16, 17$ respectively. When $\rho = 0, \ \sigma = 1$ we get

$$f'(x;0,1) = -\frac{2x(x-r_1)(x+r_1)(x^2+r_2)(x^2+r_3)}{(x-1)^2(x+1)^2(x^2+1)^2}, \quad r_1, r_2, r_3 > 0$$

which has three real roots $\{0, \pm r_1\} \in (-1,1)$ corresponding to two local maxima and one minimum of $f(x;0,1)$. By continuity it is easy to prove that the same holds for $p_\infty(x;\rho,\sigma)$. Moreover

$$\lim_{\sigma\to\infty} f'(x;\rho,\sigma) = \frac{P_2(x;\rho)}{Q_2(x;\rho)}$$

and it turns out that

$$P_2(x;\rho) = 2x(x-1)^2(x+1)^2(x^2+1)^2(x^2+2)(x\rho-1)^5$$

therefore the real roots in $(-1,1)$ are $\{0,\pm1\}$ because $1/\rho > 1$. □

We conclude that in the limit of the considered approximation the MSE is a decreasing function of n - because the support of $p_n(x)$ stretches from $I\!R$ to $(-1,1)$ as n increases - with an asymptotic value

$$MSE_\infty = \int_{-1}^{1} (x-\rho)^2 p_\infty(x) dx > 0$$

when $\sigma > 0$. Moreover the MLE are biased for all $\sigma > 0$ and MLE is a decreasing function of σ because of the continuous dependence on σ of $p_\infty(x)$ and the presence of an interval of positive probability around a secondary mode far away from ρ, which disappears only in the limit $\sigma \to 0$. In figures 1,3 simulation results are reported confirming this claims and hence the goodness of the approximation. Moreover the limit behavior for $\sigma \to \infty$ - i.e. when the data are Gaussian white noise - of $p_\infty(x)$ is the same, projected on the real line, as the condensed density [15] of the $\tilde{\zeta}_j$ obtained by solving the complex exponentials interpolation problem of a Gaussian white noise [5]. In figure 4 this behavior is illustrated by simulation. The connection between ML estimation and complex exponentials interpolation in this case justify the conjecture that the same relationship holds in the general case. As we know by [3, lemma2] that $\tilde{\gamma}_j, \tilde{\zeta}_j$ are biased we can conjecture that also the MLE in the general case are biased. Moreover we can conjecture that for σ moving from 0 to ∞ the condensed density of the MLE of ξ_j moves from a sum of Dirac's deltas centered on ξ_j to a density concentrated around the unit circle as happens for the $\tilde{\zeta}_j$.

1.4. Standard pencil methods: GPOF

Computation of MLE is usually complicated because the right model order p should be known and many local maxima are present when SNR is low or moderately large. In literature many algorithms to get approximate MLE are present and their relative merits are usually measured in terms of the CR bound for the asymptotic unbiased estimators [9, 22]. This does not make much sense because we are interested in solving the problem when $\sigma > 0$ but can help to compare algorithms. As expected because of the asymptotic unbiasdness, when the noise variance is less than a threshold, all algorithms produce reasonable estimates (see [13] for a comparison). Moreover some heuristic algorithms can exceed the CR bound (because of the bias) and hence it is suggested that the bias can help to decrease the noise threshold below which meaningful estimates can eventually be computed [22].

Moreover, because of the connection between ML estimation and complex exponential interpolation, many approximate ML algorithms are based on complex exponential interpolation of the data. The main advantages over the exact MLE algorithms are that no initialization must be provided and the computation is faster. The best of them include some sort of noise filtering in order to increase the SNR ratio. Cadzow method [10] and GPOF [18] are examples of this approach. We give here a short summary of GPOF method because it is used in the proposed estimation procedure described in Section 2 and it will be used for comparisons in Section 3. Assuming that the data $\underline{a}$ are noisy and that we know the true number p of complex exponentials, the aim of GPOF is to estimate the non linear parameters $\xi_j,\ j = 1, \ldots, p$ by solving a filtered generalized eigenvalue problem. When the data are noiseless we know that we can retrieve $\underline{\xi}$ by solving the complex exponential interpolation problem based on a square pencil of order $p \times p$ i.e. $n = 2p$ data are enough. If we use $n > 2p$ data and use a square pencil of order $n/2 \times n/2$ the conditions $detU_0 \neq 0, detU_1 \neq 0$ to solve the problem and to get a unique solution are no longer satisfied because $rank(U_0) = rank(U_1) = p < n/2$. When noise is present it make sense to assume that $n/2-p$ terms of the model represents the noise. Therefore we can solve the complex exponential interpolation problem of order

$n/2$ and then discard the $n/2-p$ terms associated e.g. with the lowest absolute values $|c_j|$ of the weights. As an alternative we can first filter-out the noise from the pencil and then solve a complex exponential interpolation problem of order p. More generally we can assume that the model is made up of l terms, $l-p$ of them representing the noise, with $p \leq l \leq n-p$, i.e. $\underline{a} = V(\underline{\tilde{\zeta}})\underline{\tilde{\gamma}}$ where $V(\underline{\tilde{\zeta}}) \in \mathbb{C}^{(n-l)\times l}$ is the Vandermonde matrix based on $\underline{\tilde{\zeta}}_j,\ j=1,\dots,l$. We notice that the larger l the smaller the number of equations $n-l$ that we can form with n observations. By choosing l we can control how accurately to represent the noise and hence the signal, but the price to pay is on the number of constraints that can be considered. We can then consider a rectangular pencil $U_1 - zU_0$ with

$$U_0 = \tilde{V}_1(\underline{\tilde{\zeta}})C\tilde{V}_2(\underline{\tilde{\zeta}})^T,\ \ U_1 = \tilde{V}_1(\underline{\tilde{\zeta}})CZ\tilde{V}_2(\underline{\tilde{\zeta}})^T$$

where $\tilde{V}_1(\underline{\tilde{\zeta}}) \in \mathbb{C}^{(n-l)\times l}, \tilde{V}_2(\underline{\tilde{\zeta}}) \in \mathbb{C}^{l\times l}$ are the Vandermonde matrices based on $\underline{\tilde{\zeta}}_j, j = 1,\dots,l$ and

$$C = diag\{\tilde{\gamma}_1,\dots,\tilde{\gamma}_l\} \text{ and } Z = diag\{\tilde{\zeta}_1,\dots,\tilde{\zeta}_l\}$$

and therefore $U_1\tilde{V}_2(\underline{\tilde{\zeta}})^{\ddagger} = U_0\tilde{V}_2(\underline{\tilde{\zeta}})^{\ddagger}Z$ where $X^{\ddagger} = (X^{\dagger})^T = (X^T)^{\dagger}$ and $X^{\dagger}$ denotes the generalized inverse of X. Therefore $\underline{\tilde{\zeta}}$ are the generalized eigenvalues of the rectangular pencil $U_1 - zU_0$. We want now to compute the signal related generalized eigenvalues by solving an eigenvalue problem of order p. To this aim let us define the data matrix

$$U = \begin{bmatrix} a_0 & a_1 & \dots & a_l \\ a_1 & a_2 & \dots & a_{l+1} \\ . & . & \dots & . \\ a_{n-l-1} & a_{n-l} & \dots & a_{n-1} \end{bmatrix} \in \mathbb{C}^{(n-l)\times(l+1)},\ p \leq l \leq n/2, \qquad (3)$$

from which we can retrieve U_0, U_1 by

$$U_0 = UE_0,\ U_1 = UE_1,\ E_0 = [\underline{e}_1,\dots,\underline{e}_l],\ E_1 = [\underline{e}_2,\dots,\underline{e}_{l+1}]$$

where $\underline{e}_j$ is the $j-$th column of the identity matrix I_{l+1}. Let us consider then its singular value decomposition $U = PDQ,\ P \in \mathbb{C}^{(n-l)\times(n-l)},\ D \in \mathbb{C}^{(n-l)\times(l+1)},\ Q \in \mathbb{C}^{(l+1)\times(l+1)}$. In the noiseless case $rank(U) = p$ therefore the last $n - l - p$ elements on the diagonal of D are zero and $U = P^F D^F Q^F$ where $D^F \in \mathbb{C}^{p\times p}$ is obtained from D by dropping the last $n - l - p$ rows or columns, $P^F \in \mathbb{C}^{(n-l)\times p}$ is obtained from P by dropping the last $n - l - p$ columns and $Q^F \in \mathbb{C}^{p\times(l+1)}$ is obtained from Q by dropping the last $n - l - p$ rows. In the noisy case we can filter out the smallest $n-l-p$ elements on the diagonal of D setting them to zero. But then the Hankel structure of $U^F = P^F D^F Q^F$ is lost. Cadzow [10] suggests to retrieve this structure while filtering out the smallest singular values by the iteration:

```
• $U^{(0)} = U$
• for $k = 0, 1, \ldots$
•       $U^{(k)} = P^{(k)} D^{(k)} Q^{(k)}$
•       $U^F = (P^{(k)})^F (D^{(k)})^F (Q^{(k)})^F$
•       $U^{(k+1)} =$ Hankel$(U^F)$
•       if $\|U^{(k+1)} - U^{(k)}\| < \eta$ then stop
• end
```

where $\eta > 0$ is a small tolerance and the operator Hankel(A) maps the matrix A into the matrix obtained by substituting each element of a secondary diagonal of A by the average of the elements of that diagonal. In [10] is proved that this iteration is a specific instance of a general method which converges under hypotheses that are verified in the case considered here. Denoting by $P^F D^F Q^F$ the singular value decomposition of the Hankel matrix produced by the iteration we are therefore reduced to solve the rectangular $(n - l) \times l$ generalized eigenvalue problem

$$P^F D^F Q^F E_1 W = P^F D^F Q^F E_0 W Z.$$

We notice that $\tilde{P} = P^F D^F \in \mathbb{C}^{(n-l)\times p}$ has maximum rank p therefore its generalized inverse is $\tilde{P}^\dagger = (\tilde{P}^H \tilde{P})^{-1} \tilde{P}^H$. Therefore by left-multiplying by $\tilde{P}^\dagger$ the problem above reduces to the rectangular $p \times l$ generalized eigenvalue problem

$$Q^F E_1 W = Q^F E_0 W Z$$

and we are looking for the p non-zero rank reducing numbers $\zeta_1^F, \ldots, \zeta_p^F$ which solve this problem. We have

Proposition 5 *Let be* $\tilde{Q}_0 = (Q^F E_0)^H \in \mathbb{C}^{l \times p}$, $\tilde{Q}_1 = (Q^F E_1)^H \in \mathbb{C}^{l \times p}$. *Then the eigenvalues of* $\tilde{Q}_0^\dagger \tilde{Q}_1 \in \mathbb{C}^{p \times p}$ *are rank reducing numbers of the rectangular pencil* $Q^F E_1 - z Q^F E_0$.

Proof. $\tilde{Q}_0$ has maximum rank p, therefore its generalized inverse is $\tilde{Q}_0^\dagger = (\tilde{Q}_0^H \tilde{Q}_0)^{-1} \tilde{Q}_0^H$ and $\tilde{Q}_0^\dagger \tilde{Q}_1$ is a square matrix of maximum rank p with non zero eigenvalues which are also the eigenvalues of $(\tilde{Q}_0^\dagger \tilde{Q}_1)^H = \tilde{Q}_1^H (\tilde{Q}_0^\dagger)^H$. But then there exist left eigenvectors $\underline{y}_j$ such that

$$\underline{y}_j^H (Q^F E_1 - \zeta_j^F Q^F E_0) = 0, \; j = 1, \ldots, p, \quad \Box.$$

We notice that the singular value decomposition of U can be replaced by its PRQ rank revealing decomposition [12] where P and Q are unitary matrices and R is a trapezoidal matrix such that the absolute values on the diagonal are in decreasing order. In fact it turns out that in the noiseless case R is a trapezoidal matrix of rank p [17, Section 7.3] and noise filtering can be performed by setting to zero the last $n-l-p$ rows of R. Despite the obvious computational advantages this method is worse than the one based on svd for low SNRs because the best approximation property of svd does not hold.

It turns out that the method is sensitive not only to the good choice of p but to the number n of data too, especially for low SNRs (see figure 2 where MSE for GPOF estimates are plotted as a function of n for moderately large noise). An easy heuristic argument can be provided for the damped sinusoids case. If k_τ is such that for $k > k_\tau$ the signal is decayed under the noise threshold (e.g $a_k \in [-3\sigma, 3\sigma], \; k > k_\tau$) more noise is injected in the estimation method hiding the signal and worsening the quality of the estimates. If n is too large it can be expected that the model order p is overestimated by any model order selection criterium such as e.g. BIC. This makes it difficult to distinguish between modes related to signal and modes related to noise. But even when n is chosen correctly, if the SNRs are moderate or small, the bias of the estimates can be the main responsible of the inability to resolve modes with close frequency. In this case the model order p is likely to be underestimated by BIC and all the estimates of the other parameters are likely to be meaningless.

2. The proposed method

2.1. Outline

From the discussion of the previous section, in order to propose an automatic method which improves on the bias affecting exact and approximate MLE, we start from the complex exponential interpolation problem, which is likely to capture the best features of MLE and exploits the ensemble behavior (as specified below) of its solution which is easier to study than the ensemble behavior of MLE. Specifically the basic observation which motivates the proposed method is the following. When SNRs are moderate or low the performances of a good standard algorithm, such as e.g. GPOF, measured by the MSE of the parameters vary significantly as a function of the noise realization used. For example for some noise realizations, two modes with close frequencies can be well separated even if SNRs are low, while for other noise realizations, with the same variance, this is not true. This means that the bias of the frequency estimates in some cases makes the two modes even closer than they are making it impossible to separate them while in other cases the opposite is true. The idea is then to base the inference on the ensemble behavior instead than on a single realization. However usually we have just one single data set. Therefore we propose to use it first to get information on the statistical distribution over the ensemble of the $\tilde{\zeta}_j, j = 1, \dots, n/2$ which are the critical quantities which the parameter estimates are based on, and then to make use of the data again to get point and interval estimates of the parameters by a stochastic perturbation method. To this purpose the original problem is reformulated as the one of estimating the complex measure

$$S(z)=\sum_{j=1}^{p} c_j\delta(z-\xi_j),\ \ \xi_j\in \text{int}(D),\ \ \xi_j\neq\xi_h\ \forall j\neq h,\ \ c_j\in\mathbb{C}$$

where $D\subset\mathbb{C}$ is a compact set, from its noisy moments

$$a_k=f(k\Delta)+\epsilon_k,\ k=0,\dots,n-1.$$

It turns out that

$$s_k=\int_D z^k S(z)dz=\iint_D (x+iy)^k S(x+iy)dxdy,\ \ k=0,1,2,\dots$$

where

$$s_k=\sum_{j=1}^{p} c_j\xi_j^k=f(k\Delta) \tag{4}$$

hence this problem is equivalent to the original one. We notice that $S(z)$ is an atomic measure supported on the (unknown) points $\xi_j,\ j=1,\dots,p$. Estimating a set Ω such that $\xi_j\in\Omega,\ j=1,\dots,p$, is our first goal.

2.2. The first step

The idea is to make use of the relation, discussed in Section 1, between the numbers $\xi_j,\ j=1,\dots,p$ and the r.v. $\tilde{\zeta}_j, j=1,\dots,n/2$ which solve the complex exponential interpolation problem for the data $a_k,\ k=0,\dots,n-1$. More specifically we want to study the location in $\mathbb{C}$ of the $\tilde{\zeta}_j$. As these are r.v. we are looking for a probability function $h(z)$ defined on the complex plane such that

$$\int_N h(z)dz=\frac{2}{n}\sum_{k=1}^{n/2}\mathcal{P}\{\tilde{\zeta}_k\in N\},\ \ N\subset\mathbb{C}.$$

The main reason to consider the $\tilde{\zeta}_j$ is now apparent: as $\tilde{\zeta}_j$ are the generalized eigenvalues of the pencil $U_1(\underline{a})-zU_0(\underline{a})$, they are the roots of the polynomial $Q(z)=$

$det(U_1(\underline{a}) - zU_0(\underline{a}))$. But then $h(z)$ is the condensed density of these roots which is given by (e.g. [5]):

$$h(z) = \frac{1}{4\pi}\Delta u(z)$$

where Δ denotes the Laplacian operator with respect to x, y if $z = x + iy$ and

$$u(z) = \frac{1}{p}E\left\{\log(|Q(z)|^2)\right\} \tag{5}$$

is the corresponding logarithmic potential and E is the expectation operator w.r.to the density of the a_k. In the limit for $\sigma \rightarrow 0$ it can be shown [3] that $h(z)$ tends weakly to a measure supported on the points $\xi_j,\ j = 1, \ldots, p$. Therefore our first goal is reached if we are able to compute the expectation in (5) and to cope with the fact that $h(z)$ conveys the information on the $\xi_j,\ j = 1, \ldots, p$ only in the limit for $\sigma \rightarrow 0$. In [4] a closed form approximation to $h(z)$ based on a single realization $\{a_k,\ k = 0, \ldots, n-1\}$ is provided. The logarithmic potential $\log(|Q(z)|^2$ is approximated by a sum of powers of χ^2 variables whose expectation can be computed in closed form. We then get

$$\hat{h}(z) \propto \sum_{k=1}^{n/2} \hat{\Delta}\left(\Psi\left[\frac{1}{2}\left(\frac{\hat{R}^2_{kk}(z)}{\sigma^2\beta} + 1\right)\right]\right) \tag{6}$$

where $\hat{\Delta}$ is the discrete Laplacian evaluated on a square lattice $\mathcal{L}$, $\hat{R}^2_{kk}(z)$ is the diagonal of the R factor in the QR factorization of $U_1 - zU_0$ and β is an hyperparameter, to be discussed in the following, which control the smoothness of $h(z)$ hence helping in coping with the noise. In fact, because of the limit property of $h(z)$, if σ is small enough there exist disjoint sets $N_k,\ k = 1, \ldots, p$, centered on $\xi_k,\ k = 1, \ldots, p$, such that $\int_N h(z)dz \approx 1, \quad N = \bigcup_k N_k$. Moreover it was shown in [5, 3] that $h(z)$ can have other noise-related local maxima which are located close to the unit circle. However if there exist signal-related local maxima close to the unit circle they can be distinguished from the noise-related ones not only by their relative higher magnitude but also by the fact that they are surrounded by a set where $h(z) \approx 0$ (gap of poles of the Pade' approximants [24, 25]). Increasing β will depress the local maxima of $h(z)$ and will make larger the sets N_k because $h(z)$ is a probability density. Eventually some sets N_k

will merge together therefore determining a loss of resolution but the local noise-related maxima will be depressed too and therefore can be easily detected and filtered out by a simple thresholding technique which can also make use of the "gap of poles" property. Furthermore only a fraction $\tilde{n} = 2\tilde{p} < n,\ \ \tilde{p} \gg p$ of data are used in this step in order to make an implicit noise filtering. Of course we loose in resolution but this is not relevant in this step. Summing up, in the first step of the procedure the data are used to identify the sets $N_k,\ k = 1, \ldots, p_N \leq p$ such that $\xi_j \in N = \bigcup_k N_k\ \forall j = 1, \ldots, p$. In fig.5 top left the results obtained at the end of the first step are shown on a specific example described in Section 4. Three not intersecting sets N_h are computed which contains in their union the true generalized eigenvalues $\xi_k, k = 1, \ldots, 5$.

2.3. The second step

Our second goal is to get point and interval estimates of the parameters. To this purpose a method based on the stochastic perturbation idea proposed in [3] is used. Pseudosamples are generated from $\{a_k,\ k = 0, \ldots, n-1\}$ by

$$a_k^{(r)} = a_k + \nu_k^{(r)},\ \ k = 0, \ldots, n-1;\quad r = 1, \ldots, T$$

where $\nu_k^{(r)}$ are i.i.d. zero mean complex Gaussian variables with variance σ'^2 independent of $a_h,\ \forall h$. The complex exponentials interpolation problem (CEIP) is solved for each of them. GPOF method is used with all data and $l = n/2,\ \hat{p} = \tilde{p}$. The generalized eigenvalues are pooled and those not belonging to N are discarded. Then a standard clustering method such as e.g. K-means [23] is applied to the generalized eigenvalues belonging to N by fixing to $\tilde{p}$ the number of cluster to be estimated and initial centroids given by the solution of the CEIP problem for the original data. The clusters whose cardinality is not close to T are discarded because it was proved in [6] that for each pseudosample it can be expected that in a small neighbor of each $\xi_k, k = 1, \ldots, p$, it will fall at least one estimated generalized eigenvalue. The number of selected clusters is an estimate $\hat{p}$ of p. In fig.5 top right and bottom left and right the big dots indicates the generalized eigenvalues which belong to N on a specific case and small dots indicates the generalized eigenvalues which do not belong to N. We notice the presence of several spurious clusters of generalized eigenvalues which justify the importance of the first step of the procedure. The estimates $\hat{\xi}_k$ of ξ_k are then computed by averaging the generalized eigenvalues belonging to the $k-$th cluster. The estimates $\hat{c}_k$ of c_k are then computed by solving the standard least squares problem

$$\underline{\hat{c}} = \text{argmin}_{\underline{\gamma}} \|V(\underline{\hat{\xi}})\underline{\gamma} - \underline{a}\|^2.$$

We notice that interval estimates of $\underline{\xi}$ and $\underline{c}$ can also be obtained from the clustering results. Remembering that we have noticed that approximate MLE can depend on the number of data n, the procedure outlined above can be improved by repeating the two steps for several values of n and choose the best one by using an order selection criterium such as e.g. BIC.

2.4. Estimation of β

The first step of the procedure depends critically on the choice of β. A value of β too small will give rise to many modes of $h(z)$ which are likely to be spurious but not easily detectable as noise-related ones. A value of β too large will give rise to a small number of modes, possibly much less than p. The clustering method can then become critical. The idea for getting a good value for β is based on a comparison of formula (6) with another approximation of $h(z)$ given in [3] by:

$$\tilde{h}(z) = \frac{1}{2\pi n}\Delta \sum_{\mu_j(z)>0} \log(\mu_j(z))$$

where $\mu_j(z)$ are the eigenvalues of

$$(U_1(\underline{s}) - zU_0(\underline{s}))\overline{(U_1(\underline{s}) - zU_0(\underline{s}))} + \frac{n\sigma^2}{2}A(z, \overline{z})$$

where $A(z,\overline{z}) \in \mathbb{C}^{n/2\times n/2}$ is a tridiagonal hermitian matrix with $1+|z|^2$ on the leading diagonal and $-\overline{z}$ and $-z$ on the diagonals respectively below and above the leading one. As the numbers $\underline{s}$ given in (4) are unknown, this formula cannot be used to estimate $h(z)$. However the following Proposition holds

Proposition 6 *If* $\beta = \frac{n}{2}$ *then* $\forall z, \quad \hat{u}(z,\sigma) - u(z,\sigma) = o(\sigma),$ *as* $\sigma \to 0$, *where* $\hat{u}(z,\sigma)$ *and* $u(z,\sigma)$ *are the logarithmic potentials respectively of* $\hat{h}(z,\sigma)$ *and* $h(z,\sigma)$

Proof. We notice that $\tilde{h}(z;\sigma)) = \frac{1}{2\pi n}\Delta \log\det(UU^H + \frac{n}{2}\sigma^2 A)$ where $U = U_1(\underline{s}) - zU_0(\underline{s})$. Let be $U = QR$ the QR decomposition of U. As $U = U^T$ we also have $U = R^T Q^T$ and therefore $UU^H = R^T Q^T \overline{QR} = R^T\overline{R}$ because Q is unitary. But then $\tilde{h}(z;\sigma)) = \frac{1}{2\pi n}\Delta \log\det(R^H R + \frac{n}{2}\sigma^2 \overline{A})$. Because of the structure of A it is easy to show that $\det(A(z)) = \frac{1-|z|^{n+2}}{1-|z|^2} \geq 1$ as $\det(A(z))$ is an increasing function of $|z|$ and $\det(A(0)) = 1$. But then, from the general identity $\det(Z+XY^H) = \det(Z)\det(I+Y^H Z^{-1}X)$ it follows that

$$
\begin{aligned}
&\det\left(R^H R + m\sigma^2\overline{A}\right) = \det\left(m\sigma^2\overline{A}\right)\det\left(I + R(m\sigma^2\overline{A})^{-1}R^H\right)\\
&\geq \det\left(m\sigma^2 I\right)\det\left(I + R(m\sigma^2 I)^{-1}R^H\right) = \det\left(RR^H + m\sigma^2 I\right)\\
&= \prod_k \lambda_k\left(RR^H + m\sigma^2 I\right) = \prod_k\left[\lambda_k(RR^H) + m\sigma^2\right] = \prod_k\left[\sigma_k(R)^2 + m\sigma^2\right]
\end{aligned}
$$

where $m = n/2, \ \ \lambda_k(X)$ denotes the $k-$th eigenvalue of X and $\sigma_k(X)$ denotes its $k-$th singular value. From [26, D.1,pg.228] we know that

$$
|R_{11},\ldots,R_{mm}| \prec_w [\sigma_1(R),\ldots,\sigma_m(R)]
$$

and, from [26, A.2,pg.116] it follows that $\forall\alpha > 0$

$$
[R_{11}^2+\alpha,\ldots,R_{mm}^2+\alpha] \prec_w [\sigma_1(R)^2+\alpha,\ldots,\sigma_m(R)^2+\alpha].
$$

But then from [26, A.2.c,pg.117] we have

$$
\prod_k\left[\sigma_k(R)^2 + m\sigma^2\right] \geq \prod_k\left[R_{kk}^2 + m\sigma^2\right] \approx \prod_k\left[E[\hat{R}_{kk}^2] + m\sigma^2\right]
$$

where E is the expectation operator, and

$$\tilde{u}(z,\sigma) = \sum_k \log\left[\sigma_k(R)^2 + m\sigma^2\right] \geq \sum_k \log\left[E[\hat{R}^2_{kk}] + m\sigma^2\right] \geq E\left[\sum_k \log\left[\hat{R}^2_{kk} + m\sigma^2\right]\right] = u(z,\sigma)$$

the last inequality follows by a generalization of Jensen inequality to matrix functions defined on the set of Hermitian matrices [26, F.2.c,pg.476, E.6,pg.467].

We notice now that [4, eq.6]

$$\begin{aligned}\hat{h}(z;\sigma,\beta) &= \frac{1}{2\pi n}\Delta \sum_{k=1}^{n/2}\left(\Psi\left[\frac{1}{2}\left(\frac{E[\hat{R}^2_{kk}]}{\sigma^2\beta}+1\right)\right]\right)\\ &\approx \frac{1}{2\pi n}\Delta \sum_{k=1}^{n/2}\left(\log\left[\frac{1}{2}\left(\frac{E[\hat{R}^2_{kk}]}{\sigma^2\beta}+1\right)\right]\right)\\ &= \frac{1}{2\pi n}\Delta \sum_{k=1}^{n/2}\log\left[E[\hat{R}^2_{kk}]+\sigma^2\beta\right]\end{aligned}$$

hence, by choosing $\beta = m$,

$$\hat{u}(z,\sigma,m) \approx \sum_k \log\left[E[\hat{R}^2_{kk}] + m\sigma^2\right].$$

The thesis follows because $\tilde{u}(z,\sigma) - u(z,\sigma) = o(\sigma)$ as proved in [3, Th.3]. □

The Proposition above suggests to use $\frac{n}{2}$ as the initial guess for β and then to increase it by a little amount to get a smoother estimate of $h(z)$ useful for estimating the set Ω in the first step. In the following the value $\beta = 0.6n$ is used.

2.5. Filtering the QR decomposition

It turns out that the first step of the procedure depends critically on the QR factorization of the matrix $U_1 - zU_0$ or, as proved in [4], on that of the matrix U defined in (3). It is therefore necessary to filter out the noise from the R factor of U. This is a very delicate task which can however successfully accomplished by taking into account the special structure of the data as follows. We notice that the real and imaginary parts of the signal $f(t) = \sum_{j=1}^{p} c_j\xi_j^t$ decay to zero exponentially. However when Gaussian noise is present the tail of the data fill a rectangular region centered on the $x-$axis of width $\approx 2\sqrt{2}\sigma$. A classic way to reduce the contribution of the noise consists therefore in applying an exponential filter to force the tail of the data to go to zero as in the noiseless case. In section 2.4 we discussed the Cadzow iteration to filter out the noise in U without destroying its Hankel structure. However, in order to further improve the estimate of R, we suggest to apply a filter also after the factorization process.

To this aim we notice first that, if $U = QR,\ Q^H Q = I,\quad R$ upper trapezoidal, the main diagonal of R can be chosen to be non-negative and monotonic decreasing. In the noiseless case the last $n-p$ rows of R must be zero, as $rank(U) = p$. It can be shown experimentally that the same behavior characterizes also the absolute value of the secondary diagonals $\{|R_{h,h+l}|,\ h = 1,\ldots,p-l\},\quad l = 0,\ldots,p-1$. Moreover this behavior is preserved also in the noisy case but with an asymptotic value greater than zero. In fig.6 the results of a simulation showing these facts are reported. A set of complex exponential signals were generated with random frequencies ω_j and phases θ_j with uniform distribution in $[-\pi,\pi)$, random decays ρ_j with uniform distribution in $(0,1]$ and complex standard Gaussian random amplitudes normalized in order to make their absolute values to sum to one. The matrix U was then formed and the QR decomposition was computed. The absolute values of the diagonals of R were then averaged and the results for the main diagonal and the first three secondary diagonals was plotted. The same is done by adding complex Gaussian white noise to the complex exponential signals.

The comparison of the results in the noiseless and noisy cases for several SNRs and orders p, suggests that we can filter out the noise in the diagonals of R by

$$\tilde{R}_{h,h+l} = \frac{R_{h,h+l}}{h^{\gamma_l}},\ \ h = 1,\ldots,p-l,\ \ \gamma_l > 0,\ \ l = 0,\ldots,p-1.$$

In fig(6) the filtered diagonals were plotted too where γ was estimated by solving the problems

$$\hat{\gamma}_l = \mathrm{argmin}_{\gamma} \sum_{h=1}^{p-l} |\tilde{R}_{h,h+l} - R_{h,h+l}|,\ \ l = 0,\ldots,p-1.$$

It can be noticed a good agreement between the noiseless and filtered data, therefore the functional form of the filter seems to be adequate to do the job. In the following we choose only one hyperparameter γ and filter the diagonals of R according to the rule

$$\tilde{R}_{h,h+l} = \frac{R_{h,h+l}}{h^{\gamma}} \tag{7}$$

2.6. The algorithm

Summing up, a sketch of the proposed algorithm is the following:

- fix an overestimate $\tilde{p} \gg p$ of the true number of components p
- compute U based on the first $\tilde{n} = 2\tilde{p}$ data and filter it by Cadzow algorithm
- compute $U = QR$ and filter the diagonals of R by formula (7)
- compute the Hessemberg matrices $R(E_1 - zE_0), \ \forall z \in \mathcal{L}$ and reduce them to triangular form by Givens rotations
- compute $\hat{h}(z;\beta), \ \beta = 0.6n$, by formula (6) where $\hat{R}_{kk}(z)$ are the diagonal elements of the triangular matrices computed in the previous step
- compute the sets N_k such that
  - $\hat{h}(z;\beta)$ is unimodal for $z \in N_k$
  - $\bigcap_{k=1}^{p} N_k = \emptyset$

  by selecting the local maxima of $\hat{h}(z;\beta)$ above a given threshold $\tau > 0$, and then by identifying the neighbor N_k of the k-th local maxima $\hat{\xi}_k$ such that $\hat{h}(z;\beta)$ is monotonic decreasing along the four coordinate directions on the lattice $\mathcal{L}$ starting from $\hat{\xi}_k$
- generate T pseudosamples based on the original n data
- solve the CEIP for each pseudosample by GPOF method with parameters $l = n/2, \tilde{p}$ and pool the $\xi_h^{(r)}$
- cluster the $\xi_h^{(r)} \in \bigcup N_k$ and discard the others. The k-means method is used with $\tilde{p}$ as the number of clusters, and the clusters with less than $\lfloor \alpha T \rfloor, \quad \alpha \in (0.5, 1]$

elements are discarded

- p_{ott} = number of selected clusters
- $\hat{\xi}_k$ = average of the $\xi_h^{(r)}$ in cluster k-th, $k = 1, \ldots, p_{ott}$
- $\hat{c}_k$ = average of the $c_h^{(r)}$ in cluster k-th, $k = 1, \ldots, p_{ott}$

3. Simulation results

In order to test the advantages of the proposed method w.r.to the standard ones, four experiments were performed corresponding to the four values of the noise s.d. $\sigma = 2\sqrt{2}, \sqrt{2}, \frac{\sqrt{2}}{3}, \frac{\sqrt{2}}{10}$. In each experiment $N = 100$ independent realizations of the r.v. $a_k^{(h)}, k = 1, \ldots, n = 120, \quad h = 1, \ldots, N$ were generated from the complex exponentials model with $p = 5$ components given by

$$\underline{\xi} = \left[e^{-0.3-i2\pi 0.35}, e^{-0.1-i2\pi 0.3}, e^{-0.05-i2\pi 0.28}, e^{-0.0001+i2\pi 0.2}, e^{-0.0001+i2\pi 0.21}\right]$$

$$\underline{c} = [20, 6, 3, 1, 1]$$

by adding complex Gaussian noise with s.d. σ. We notice that the frequencies of the 4^{rd} and 5^{th} components are closer than the Nyquist frequency if $n < 1/(0.21 - 0.20) = 100$. By defining $SNR_i = \sqrt{2}\frac{|c_i|}{\sigma}$ we label the four considered cases by $SNR = \min_i SNR_i = [0.5, 1, 3, 10]$. For each experiment and for each $h = 1, \ldots, N$ the method GPOF [18] was applied with $l = m/2, \quad m = n/3, \ldots, n$ and $\hat{p} = l/3, \ldots, l/2$. For each estimate $\tilde{\underline{\zeta}}(m, \hat{p})$ of the generalized eigenvalues, the corresponding estimates $\gamma(m, \hat{p})$ of the weights was obtained by solving a linear least squares problem. The optimal values (m_{ott}, p_{ott}) of $(m, \hat{p})$ were chosen by minimizing the BIC criterium [2]. The corresponding optimal parameters $\hat{\xi}_j$ and $\hat{c}_j$ were then computed. $|\hat{c}_j|$ were then sorted in descending order and $\hat{\xi}_j$ were sorted accordingly. $\hat{c}_j$ and $\hat{\xi}_j$ were then used to estimate the signal by

$$s_k = \sum_{j=1}^{p_{ott}} \hat{c}_j \hat{\xi}_j^k.$$

If $p_{ott} \geq p$, the relative error was computed by

$$E(\sigma, h) = \frac{\sum_{j=1}^{p} |c_j - \hat{c}_j|^2}{\sum_{j=1}^{p} |c_j|^2} + \frac{\sum_{j=1}^{p} |\xi_j - \hat{\xi}_j|^2}{\sum_{j=1}^{p} |\xi_j|^2}.$$

Otherwise $E(\sigma, h)$ was set to the conventional value -1. In fig. 7 the empirical distributions over the N replications of $p_{ott}(\sigma, \cdot)$ and $E(\sigma, \cdot)$ were reported for each value of σ. It can be noted that the optimal model order is reasonably concentrated around the true value $p = 5$ only for the two smallest value of σ. Consequently also the relative error is reasonably small only in those cases. The average relative MSEs

$$MSE(\sigma) = \frac{1}{N_\sigma} \sum_{h=1}^{N_\sigma} E(\sigma, h)$$

where N_σ is the cardinality of the set $\{h | E(\sigma, h) \geq 0\}$, are reported in the first row of Table 1. In the second row the cardinalities N_σ are reported.

	$SNR = 0.5$	$SNR = 1$	SNR=3	SNR=10
$MSE(standard)$	1.4	1.2	0.3	0.1
N_σ	18	14	68	100
$MSE(proposed)$	1.2	0.9	0.4	0.1
N_σ	62	83	95	90

Table 1. Standard method: relative MSEs (first row) averaged over N_σ (second row) replications. Proposed method: relative MSEs (third row) averaged over N_σ (fourth row) replications.

The new method was then applied to the same data. The algorithm illustrated in section 2.6 for $\sigma = 2\sqrt{2}, \sqrt{2}, \frac{\sqrt{2}}{3}, \frac{\sqrt{2}}{10}$ and $h = 1, \ldots, N$ was applied. However in this case only the first $n = 80$ data values were used, hence a super-resolution problem is involved in this case. After some trials and errors the following hyperparameters provide the best results: lattice dimension = 80, number of iterations of the Cadzow algorithm = 10, filter parameter $\gamma = 0.4$, threshold for selecting the local maxima of the condensed density $\tau = 2.e - 3$, number of clusters for k-means $\tilde{p} = 20$, number of pseudosamples $T = 20$, ratio between standard deviation of pseudosamples and noise standard deviation $\frac{\sigma'}{\sigma} = 0.15$, acceptation threshold for clusters $\alpha = 0.75$. Results are reported in fig. 8. However we remark that the hyperparameters values are not critical in the sense that the same qualitative results are obtained by slightly perturbing these values.

The average relative MSEs and the corresponding cardinalities N_σ are reported in the third and fourth rows of Table 1.

The following advantages over the standard method can be noticed:

- only 2/3 of data are needed to get the best results
- the right order p is better estimated for low SNRs
- the MSE is comparable for large SNRs and better for low SNRs

We remark that the results reported in fig.8 and Table 1 were obtained by using the same hyperparameters reported above for all SNRs considered. A fine tuning of the hyperparameters could possibly further improve the results. Finally we notice that the standard method based on GPOF requires about six times more floating point operations than the proposed method in the case considered. This is due to the fact that $O(n^2/6)$ svd factorizations are required in the standard method to select the good order, while only T svd factorizations are needed in the proposed method.

4. Conclusions

A classic approximation problem which is at the core of many ill posed inverse problems arising in many application fields is revisited and a new stochastic approach is considered to overcome the drawbacks of standard methods. It turns out that some tools developed in the framework of the theory of random matrices, such as the condensed density of the (generalized) eigenvalues, provides a deep insight on the structure of the approximation problem. Coupling this information with a stochastic perturbation approach, the bias which affects standard estimators based on Maximum Likelihood can be controlled and a solution with better statistical properties, than those provided by standard methods, can be computed. The price to pay is that a few hyperparameters must be identified in order to get an automatic estimation procedure. However the flexibility provided by the hyperparameters allows to extract the correct information in low SNR situations.

References

[1] Abbey C K, Clarkson E, Barrett H H, Muller S P, Rybicki F J 1998 A method for approximating the density of maximum likelihood and maximum a posteriori estimates under a Gaussian noise model *Medical Image Analysis* **2** 395-403

[2] Akaike H 1979 A Bayesian extension of the minimum AIC procedure of autoregressive model fitting *Biometrika* **66** 237-242

[3] Barone P 2008 A new transform for solving the noisy complex exponentials approximation problem *J. Approx. Theory* **155** 1-27

[4] Barone P 2008 On the condensed density of the generalized eigenvalues of pencils of Hankel Gaussian random matrices and applications *arXiv:0801.3352*.

[5] Barone P 2005 On the distribution of poles of Padé approximants to the Z-transform of complex Gaussian white noise *J. Approx. Theory* **132** 224-240

[6] Barone P, March R 1998 Some properties of the asymptotic location of poles of Padé approximants to noisy rational functions, relevant for modal analysis. *IEEE Trans. Signal Process.* **46** 2448-2457

[7] Barone P March R 2001 A novel class of Padé based method in spectral analysis *J. Comput. Methods Sci. Eng.* **1** 185-211

[8] Barone P, Ramponi A, Sebastiani G 2001 On the numerical inversion of the Laplace transform for Nuclear Magnetic Resonance relaxometry. *Inverse Problems* **17** 77-94

[9] Bresler Y and Macovski A 1986 Exact Maximum Likelihood parameter estimation of superimposedexponential signals in noise *IEEE Trans.Ac.Sp.Sign.Proc.* **34** 1081-1089

[10] Cadzow J A 1988 Signal enhancement-a composite property mapping algorithm *IEEE Trans.Ac.Sp.Sign.Proc.* **36** 49-62

[11] Cedilnik A, Kosmelj K, Blejec A 2004 The distribution of ratio of jointly normal variables *Metodoloski zvezki* **1** 99-108

[12] Chan T F and Hansen P C 1992 Some applications of the rank revealing QR factorization *SIAM*

J. Sci. and Stat. Comput. **13** 727-741

[13] Elad M, Milanfar P, Golub G H 2004 Shape from moments-an estimation theory perspective *IEEE Trans.Sign.Proc.* **52** 1814-1829

[14] Golub G and Pereyra V 2003 Separable nonlinear least squares: the variable projection method and its applications *Inverse Problems* **19** R1-R26

[15] Hammersley J M 1956 The zeros of a random polynomial *Proc. Berkely Symp. Math. Stat. Probability 3rd* **2** 89-111.

[16] Henrici P 1977 *Applied and computational complex analysis vol.I* (New York: John Wiley)

[17] Horn R A and Johnson C R 1985 *Matrix Analysis* (Cambridge: Cam. Univ. Press)

[18] Hua Y and Sarkar T K 1989 Generalized pencil-of-function method for extracting poles of an EM system from its transient response *IEEE Trans. Antennas Propagat.* **37** 229-234

[19] Hua Y and Sarkar T K 1990 Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise *IEEE Trans.Ac.Sp.Sign.Proc.* **38** 814-824

[20] Hua Y and Sarkar T K 1991 On SVD for estimating generalized eigenvalues of singular matrix pencil in noise *IEEE Trans.Ac.Sp.Sign.Proc.* **39** 892-900

[21] Joshi M 1976 On the attainment of the Cramer-Rao lower bound *The Annals of Stat.* **4** 998-1002

[22] Kay S M 1984 Accurate frequency estimation at low signal-to-noise ratio *IEEE Trans.Ac.Sp.Sign.Proc.* **32** 540-547

[23] MacQueen J B 1967 Some Methods for classification and Analysis of Multivariate Observations *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability"* (Berkeley: University of California Press) 291-297

[24] March R, Barone P 1998 Application of the Padé method to solve the noisy trigonometric moment problem: some initial results *SIAM J. Appl. Math.* **58** 324-343

[25] March R, Barone P 2000 Reconstruction of a piecewise constant function from noisy Fourier coefficients by Padé method. *SIAM J. Appl. Math.* **60** 1137-1156

[26] Marshall A W and Olkin I 1979 *Inequalities: theory of majorization and its applications* (New York: Academic Press)

[27] Scharf L L 1991 *Statistical signal processing* (Reading: Addison-Wesley)

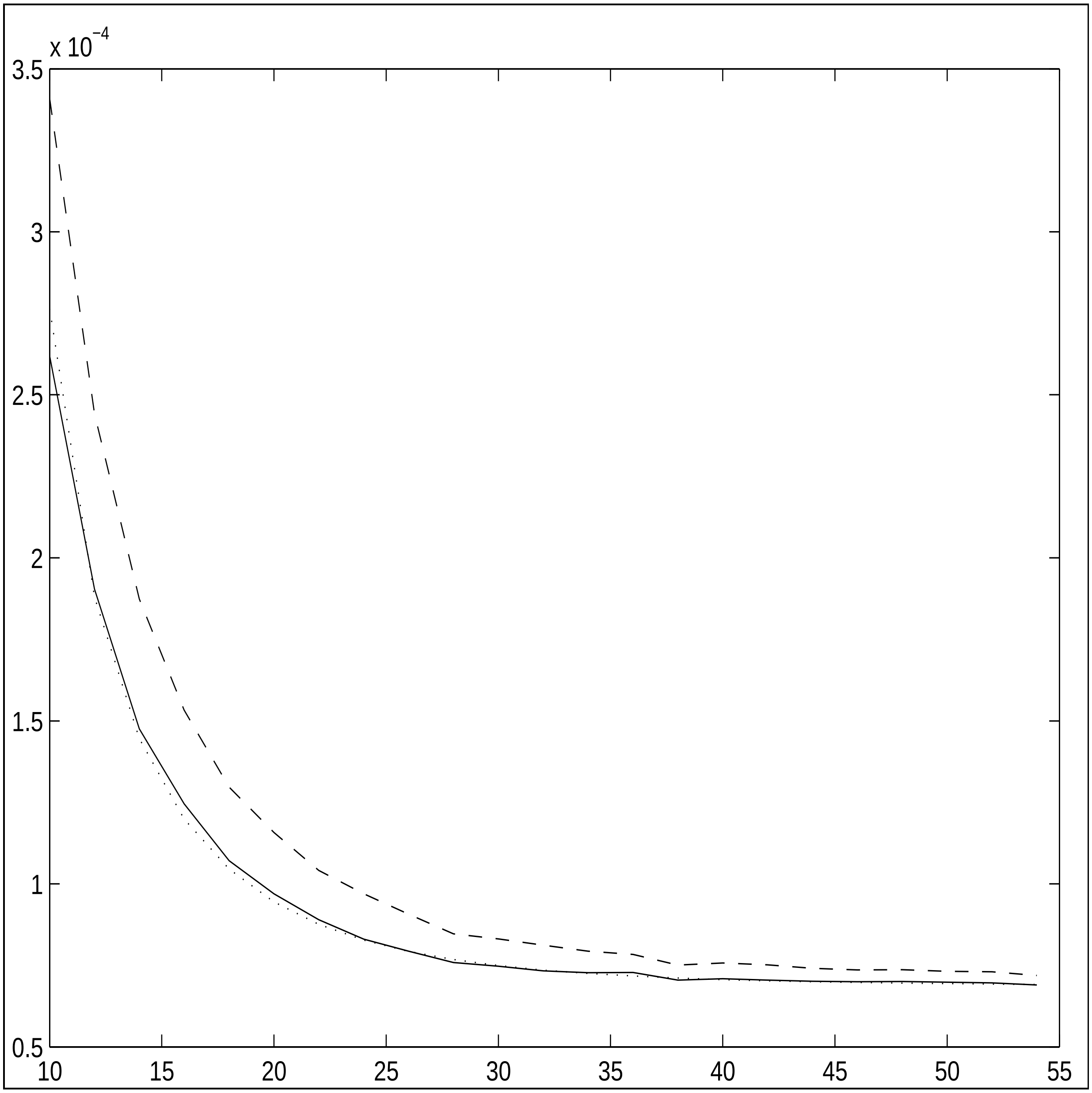

Figure 1.

MSE as a function of the number of data. MSE of $\rho_{ML} = -0.9$ of the example in Section 1 computed by MonteCarlo simulation (10000 samples, $\sigma = 0.1$) and numerical minimization (solid), MSE computed by numerical integration using the approximated density (dotted), MSE computed by MonteCarlo simulation and GPOF method with filter order 1 (dashed).

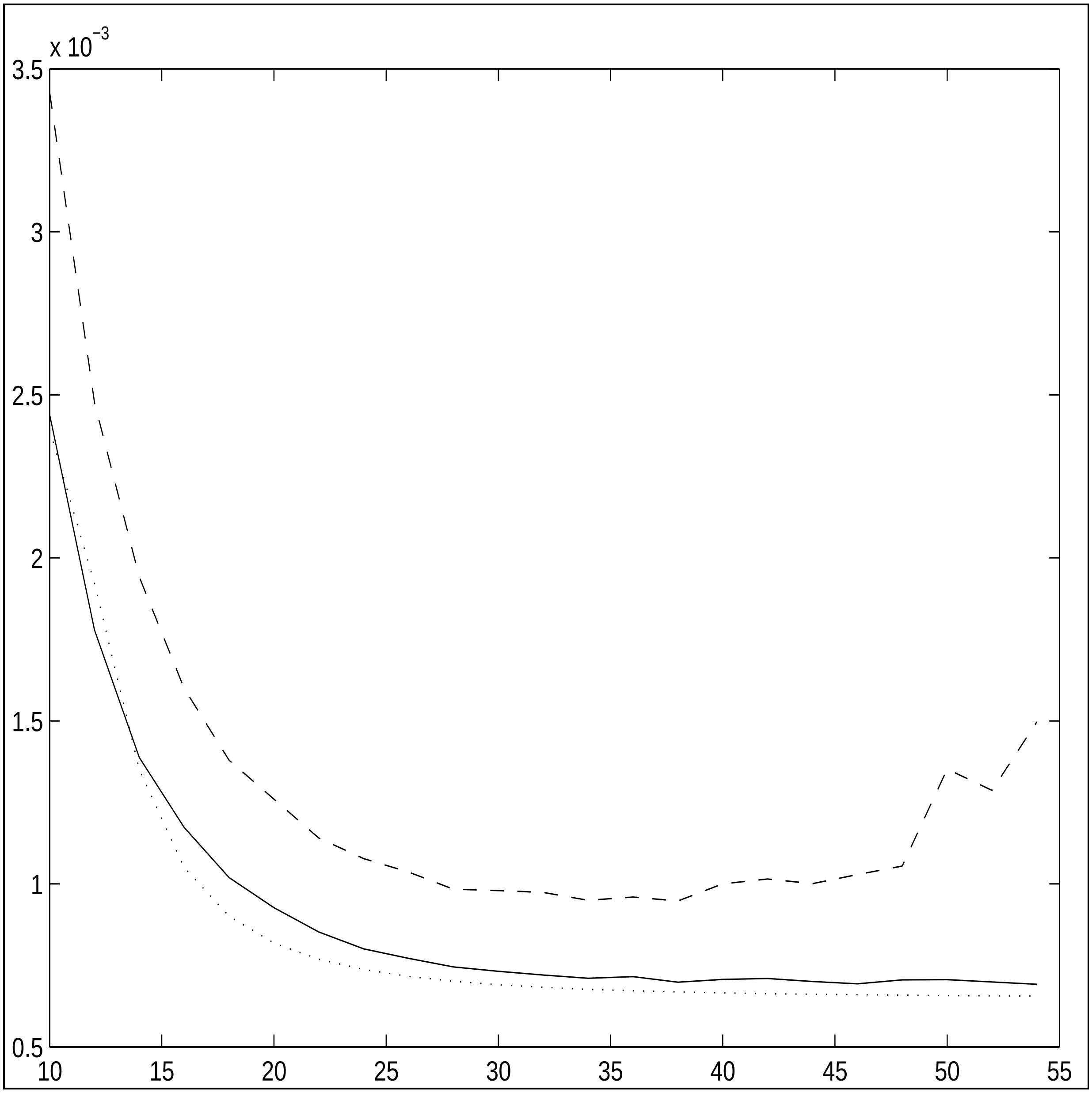

Figure 2.

MSE as a function of the number of data. MSE of $\rho_{ML} = -0.9$ of the example in Section 1 computed by MonteCarlo simulation (10000 samples, $\sigma = 0.3$) and numerical minimization (solid), MSE computed by numerical integration using the approximated density (dotted), MSE computed by MonteCarlo simulation and GPOF method with filter order 1 (dashed).

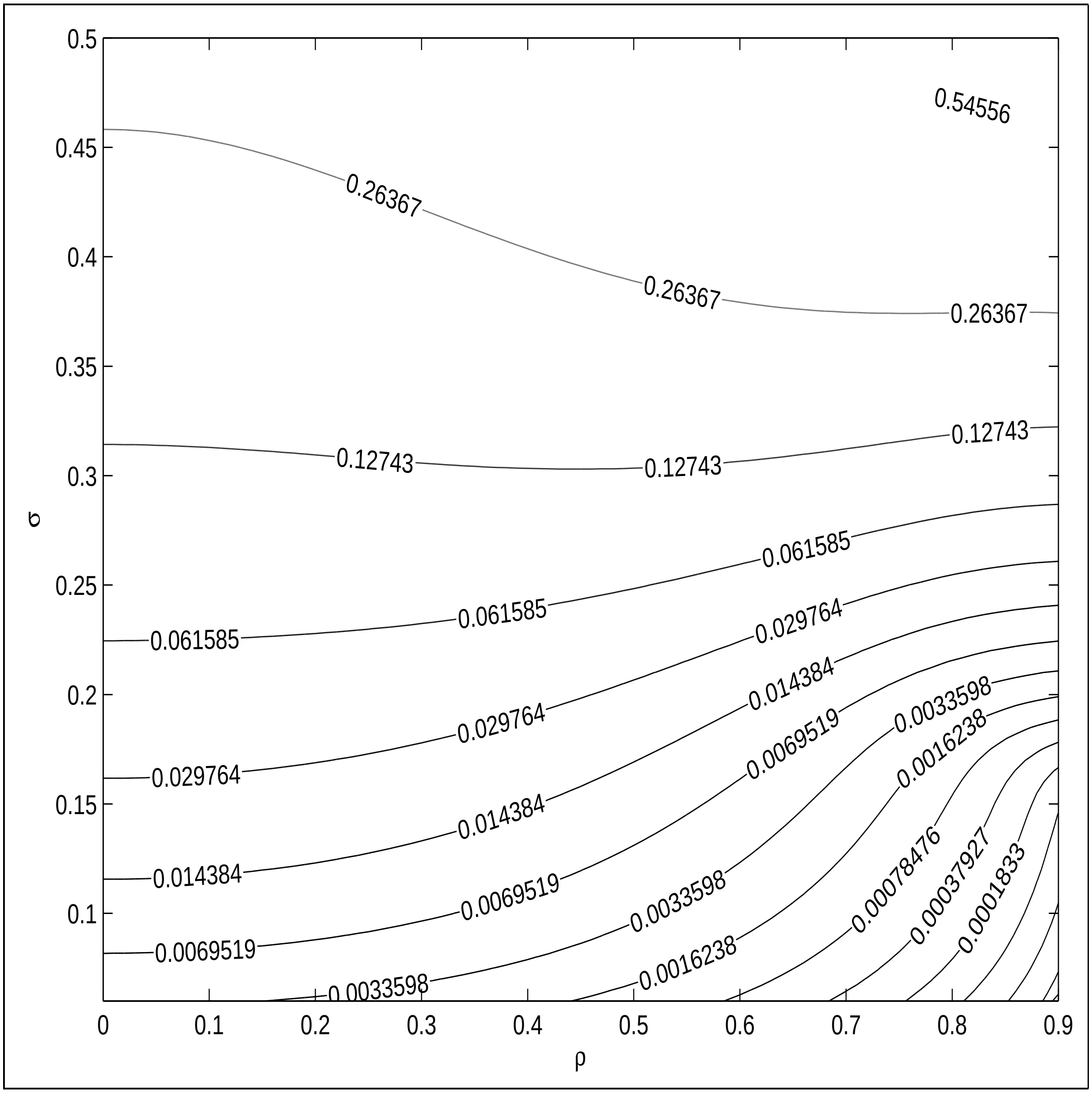

Figure 3.

MSE of ρ_{ML} of the example in Section 1 as a function of the true value of the parameter and of the noise standard deviation, computed by numerical integration using the approximated density.

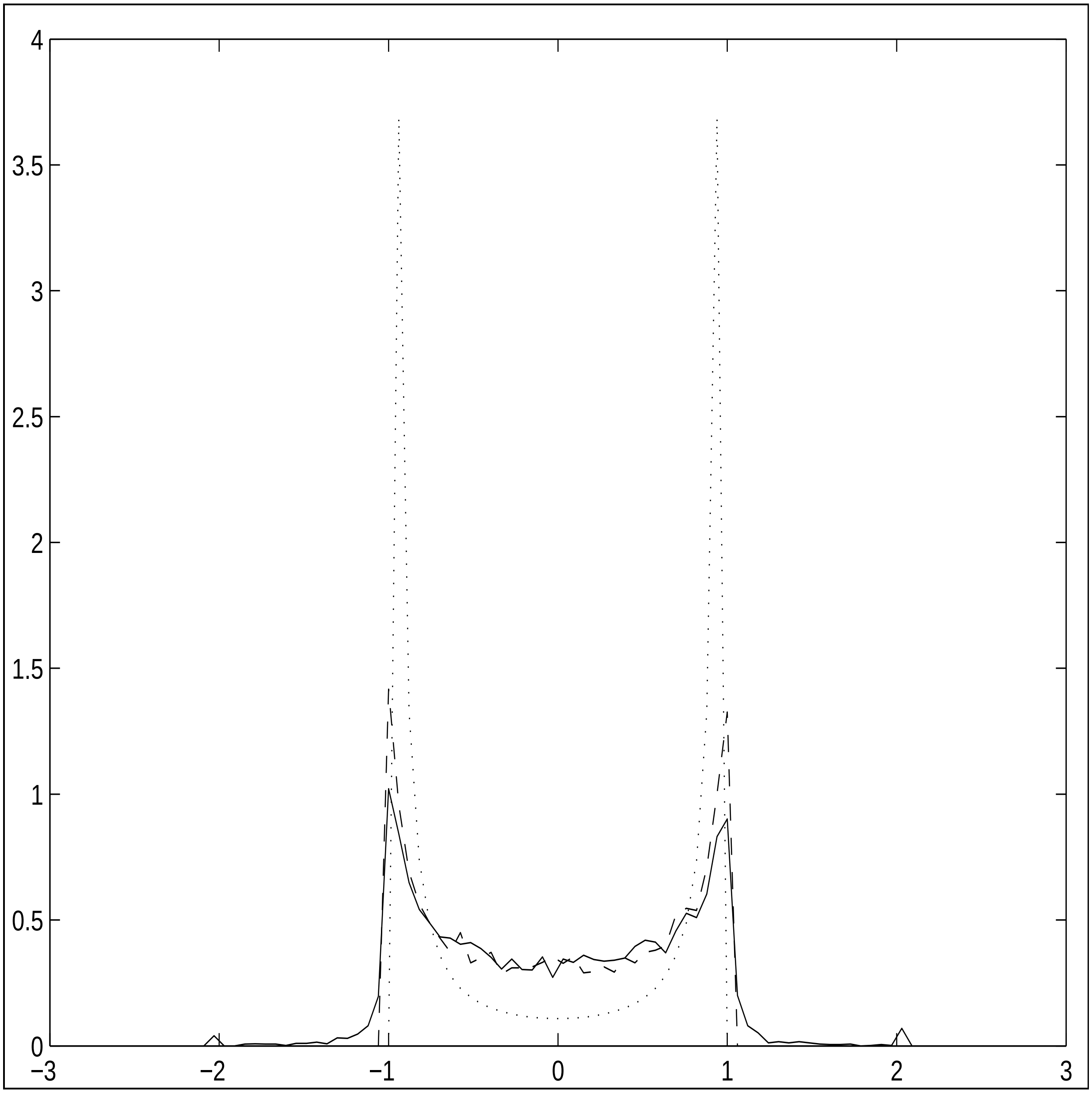

Figure 4.

Density $p_\infty(x)$ of ρ_{ML} of the example in Section 1 (dotted). Empirical density computed by MonteCarlo simulation (10000 samples, $\sigma = 100, \rho = -0.8, n = 500$) and numerical minimization with uniformly distributed initial value in $[-1, 1]$ (solid). Empirical density computed by MonteCarlo simulation and GPOF method with filter order 1 (dashed).

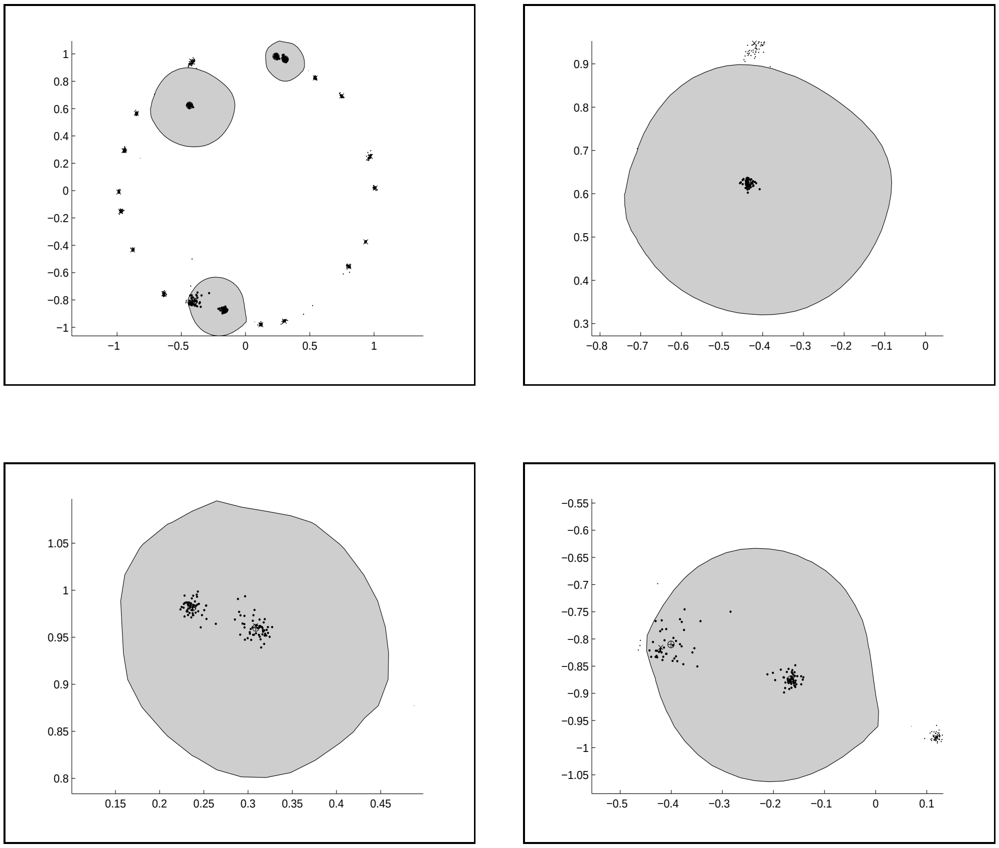

Figure 5. Top left: the sets $N_j,\ j = 1, \ldots, p_N,\ p_N = 3,$ SNR $= 0.5$; top right and bottom left and right: zoom of the sets N_1, N_2, N_3; the small dots are the generalized eigenvalues corresponding to each pseudosample; the big dots are the generalized eigenvalues falling in $N_1 \cup N_2 \cup N_3$; the ”x” are the initial centroids of the clustering procedure; the ”+” are the estimated centroids; the ”o” are the centroids of clusters with more than $0.75 \cdot R$ points where $R = 100$ is the number of psudosamples.

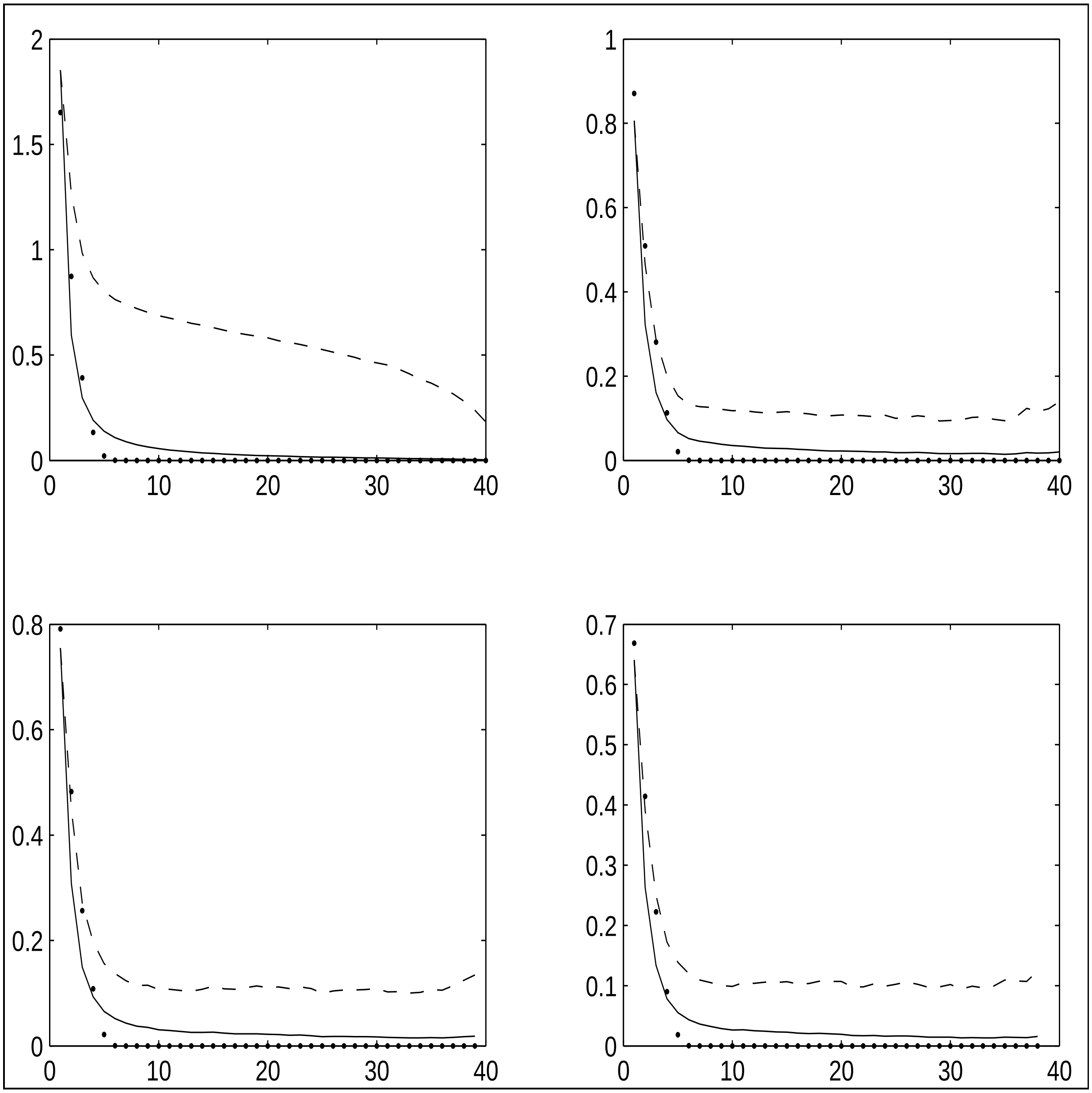

Figure 6.

Top right: the main diagonal of the matrix R in the noiseless case (dotted), in the noisy case (dashed) and the filtered one (solid) are represented when SNR$=1, \quad p=5$. Top left: the same for the absolute value of the first diagonal. Bottom left: the same for the absolute value of the second diagonal. Bottom right: the same for the absolute value of the third diagonal.

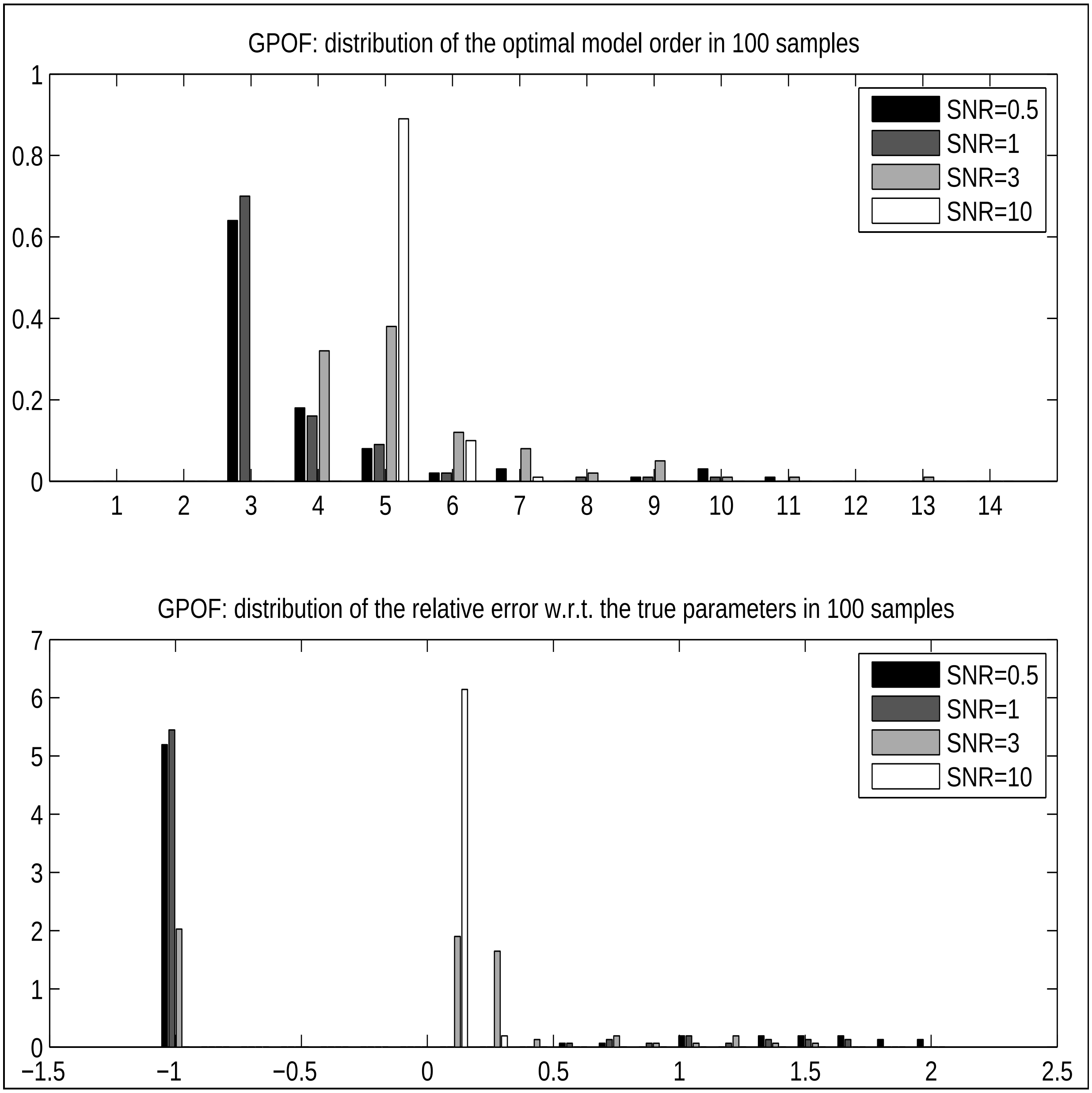

Figure 7. The standard method: the empirical distributions over the N replications of $p_{ott}(\sigma,\cdot)$ (top) and $E(\sigma,\cdot)$ (bottom) for $\sigma = 2\sqrt{2}, \sqrt{2}, \frac{\sqrt{2}}{3}, \frac{\sqrt{2}}{10}$.

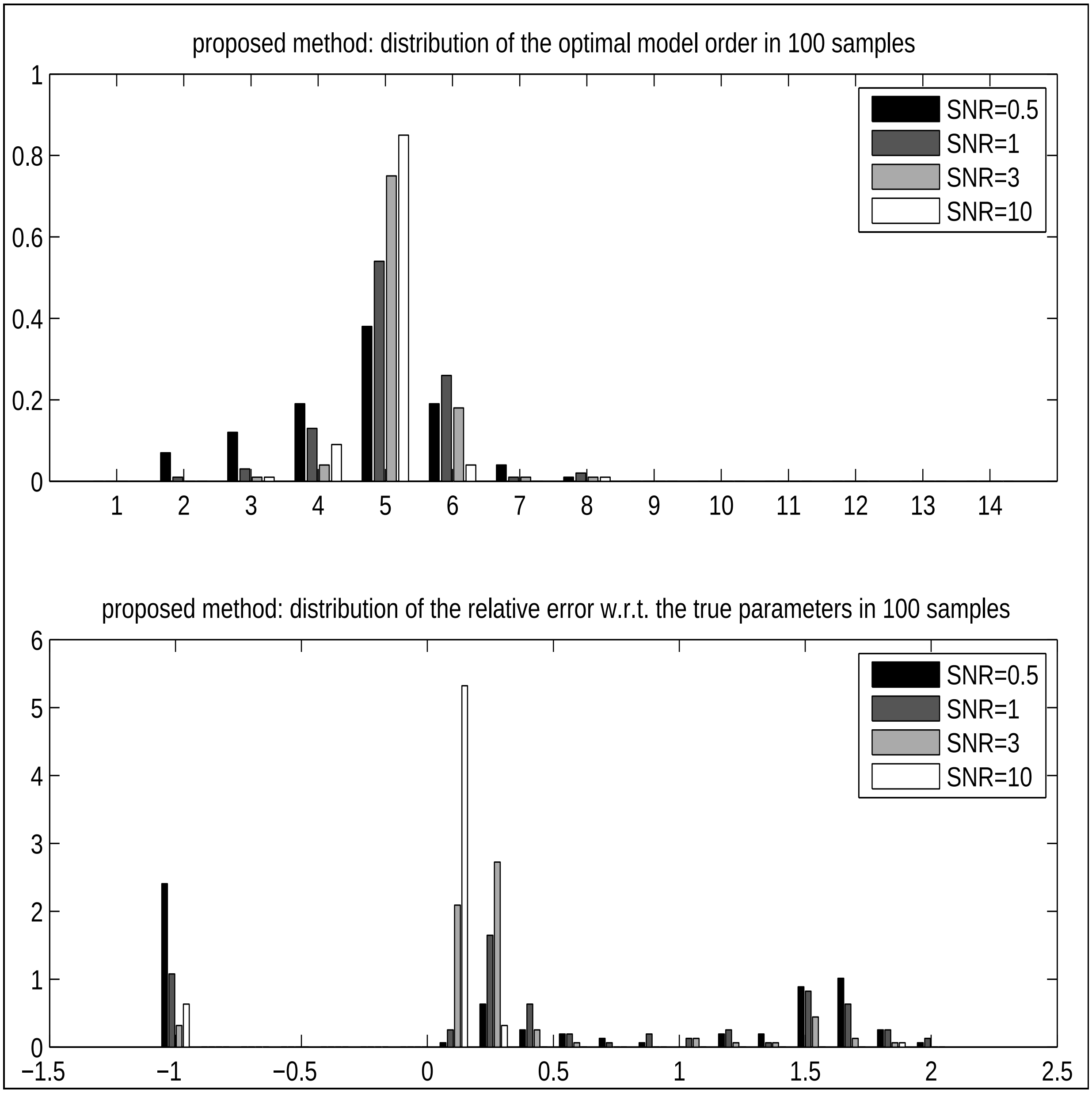

Figure 8. The proposed method: the empirical distributions over the N replications of $p_{ott}(\sigma, \cdot)$ (top) and $E(\sigma, \cdot)$ (bottom) for $\sigma = 2\sqrt{2}, \sqrt{2}, \frac{\sqrt{2}}{3}, \frac{\sqrt{2}}{10}$.